

# Water Density Fluctuations Relevant to Hydrophobic Hydration are Unaltered by Attractions

Richard C. Remsing<sup>1</sup> and Amish J. Patel<sup>1,\*</sup>

<sup>1</sup>*Department of Chemical and Biomolecular Engineering,  
University of Pennsylvania, Philadelphia, PA 19104*

(Dated: February 19, 2015)

An understanding of density fluctuations in bulk water has made significant contributions to our understanding of the hydration and interactions of idealized, purely repulsive hydrophobic solutes. To similarly inform the hydration of realistic hydrophobic solutes that have dispersive interactions with water, here we characterize water density fluctuations in the presence of attractive fields that correspond to solute-water attractions. We find that when the attractive field acts only in the solute hydration shell, but not in the solute core, it does not significantly alter water density fluctuations in the solute core region. We further find that for a wide range of solute sizes and attraction strengths, the free energetics of turning on the attractive fields in bulk water are accurately captured by linear response theory. Our results also suggest strategies for more efficiently estimating hydration free energies of realistic solutes in bulk water and at interfaces.

## I. INTRODUCTION

Hydrophobic effects and solvent-mediated phenomena in general, are important in a wide variety of contexts [1–4] ranging from protein folding [5–7] and aggregation [8–10], to colloidal assembly [11–13] and detergency [14, 15]. When a hydrophobic solute is solvated by water, it excludes water molecules from the region that it occupies, thereby perturbing the inherent interactions between the nearby water molecules [1–4, 16–18]. Because water molecules have strong hydrogen bond interactions, it is the penalty for this disruption of the water structure that dominates hydrophobic hydration free energies. Hard sphere solutes that simply exclude water molecules from the region they occupy have thus served as idealized hydrophobic solutes, and their hydration and assembly have been extensively studied using both molecular simulations [2, 19–29] and theory [3, 17, 20, 30–36].

The use of hard solutes as ideal hydrophobes is particularly judicious, because their excess hydration free energy,  $\Delta G_{\text{HS}}$ , is intimately tied to fluctuations in water density through [32, 37, 38]

$$\beta \Delta G_{\text{HS}} = -\ln P_v(N=0), \quad (1)$$

where  $P_v(N)$  is the probability of observing  $N$  waters in the volume  $v$  corresponding to the size and shape of the hard solute, and  $\beta = 1/k_B T$  ( $k_B$  is Boltzmann’s constant and  $T$  is the temperature). Thus, an understanding of density fluctuations in bulk water has the potential to inform free energies of hydrophobic hydration and association. Indeed, the notion that small fluctuations in water density obey Gaussian statistics lies at the heart of the Pratt-Chandler theory of hydrophobic hydration [30, 31]. Using molecular simulations, Hummer *et al.* explicitly showed that density fluctuations in small observation volumes ( $v \lesssim 1 \text{ nm}^3$ ) are Gaussian, allowing them to relate

$\Delta G_{\text{HS}}$  to the moments of  $P_v(N)$  [32, 33]. This simplification has afforded molecular insights into various phenomena, including entropy convergence, wherein protein unfolding entropies converge at a common temperature [28, 33], as well as the denaturation of proteins [39–41] and the formation of clathrate hydrates [41] at high pressures.

Noting that water at ambient conditions is in close proximity to liquid-vapor coexistence, the theory of Lum, Chandler, and Weeks (LCW) predicted [34, 36, 42] that low- $N$  fluctuations should be enhanced (relative to Gaussian statistics) in large volumes ( $v \gtrsim 1 \text{ nm}^3$ ). Indeed, simulations subsequently verified that while  $P_v(N)$  remains Gaussian near its mean, the likelihood of low- $N$  fluctuations in large  $v$  is enhanced significantly in bulk water and even more so near hydrophobic surfaces; that is,  $P_v(N)$  develops fat low- $N$  tails [4, 26, 27, 42–48]. This perspective clarified that water near hydrophobic surfaces sits at the edge of a dewetting transition that can be readily triggered by small perturbations [48–53]. It also led to the prediction that the assembly of small hydrophobic solutes in the vicinity of extended hydrophobic surfaces would be barrierless [47, 54]. Thus, an understanding of density fluctuations in small and large volumes as well as in the vicinity of interfaces has made significant contributions to our understanding of hydrophobic hydration and assembly.

In addition to excluding water, realistic solutes also possess favorable attractive (van der Waals) interactions with water. Here, our goal is to establish a connection between water density fluctuations and hydration free energies of *realistic* solutes of various sizes in bulk water and at interfaces. To accomplish this, we first turn on solute-water attractive interactions, and then characterize water density fluctuations corresponding to the emptying of the solute repulsive core in the presence of these attractions. The hydration free energy of realistic van der Waals solutes,  $\Delta G_{\text{vdW}}$ , is then given by:

$$\beta \Delta G_{\text{vdW}} = \beta \Delta G_{\text{att}} - \ln P_v^{(\text{att})}(N=0), \quad (2)$$

\* amish.patel@seas.upenn.edu

where  $\Delta G_{\text{att}}$  is the free energy for turning on the solute-water attractions in bulk water, and  $P_v^{(\text{att})}(N)$  is now the probability of observing  $N$  waters in  $v$  in the presence of the attractive field. We show that  $\Delta G_{\text{att}}$  is accurately described by linear response theory, so that an understanding of  $P_v^{(\text{att})}(N)$  informs  $\Delta G_{\text{vdW}}$  in much the same way as  $P_v(N)$  has informed  $\Delta G_{\text{HS}}$ .

We quantify  $P_v^{(\text{att})}(N)$  for spherical volumes of different sizes and a range of attractive strengths. We find that the presence of attractions in both the solute core,  $v$ , and its hydration shell, following the Weeks-Chandler-Andersen (WCA) prescription [55], significantly alters water density fluctuations; it becomes progressively harder to empty  $v$  as the strength of attractions is increased. However, as far as the estimation of  $\Delta G_{\text{vdW}}$  is concerned, this relative difficulty in emptying  $v$  in the presence of attractions is largely offset by the favorable free energy,  $\Delta G_{\text{att}}$ , of turning on those attractions in the first place.

To minimize cancellation between the favorable  $\beta\Delta G_{\text{att}}$  and the unfavorable  $-\ln P_v^{(\text{att})}(N=0)$  terms in Equation 2, we consider an alternative prescription for hydrating the same solute; one that involves turning on attractions in the hydration shell, but not in the solute core,  $v$ . The overall value of  $\Delta G_{\text{vdW}}$  does not depend on which prescription is used and in particular, whether attractions are turned on in the solute core or not; attractions in the core simply increase the magnitude of the components of  $\Delta G_{\text{vdW}}$ . In the presence of attractions in the hydration shell alone, we find that the water density fluctuations in the core are remarkably unaltered; attractions effect only a subtle change in the mean density. We find this to be true for solutes of all sizes and reasonable attraction strengths. Thus, our primary finding is that water density fluctuations that are relevant to hydrophobic hydration are largely unaffected by the presence of attractions.

Our results also suggest strategies for circumventing the breakdown of perturbation theories of hydrophobic hydration, which occurs for large ( $\gtrsim 1 \text{ nm}^3$ ) solutes [51, 56, 57].  $\Delta G_{\text{vdW}}$  is typically estimated by first creating a cavity and subsequently turning on attractions. The free energy for turning on attractions can be readily estimated if accurate approximations are available for the response of the hydration shell water density to the attractions. This approach works well for small solutes because water structure around the cavity is not significantly altered when attractions are turned on, so that water density responds linearly to attractions. In contrast, a soft vapor-liquid interface is nucleated around large cavities, and is readily displaced towards the solute when attractions are turned on. As a result, water density near the cavity increases in a sigmoidal fashion (and not linearly) as the strength of attractions is increased, thereby violating a key assumption that underpins perturbation theory. Because we turn on attractions in bulk water prior to the formation of a cavity, water density responds linearly to the strength of attractions, and  $\Delta G_{\text{att}}$

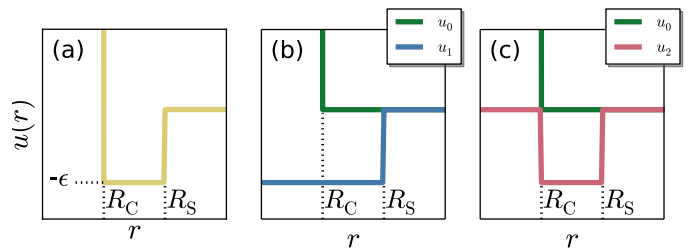


FIG. 1: (a) The square well potential,  $u(r)$ , with a hard core repulsion at  $r = R_C$  and a well of depth,  $\epsilon$ , and width,  $R_S - R_C$ . (b) This solute-water potential can readily be split into its repulsive,  $u_0(r)$ , and attractive,  $u_1(r)$ , components using a WCA-like prescription [55]. (c) Alternatively,  $u(r)$  can be split into  $u_0(r)$ , and the attractive potential,  $u_2(r)$ , which is non-zero only in the shell region,  $R_C < r < R_S$ .

is readily estimated using linear response theory. Water density fluctuations are largely unaltered when attractions are turned on in the hydration shell alone, so that  $-\ln P_v^{(\text{att})}(N=0)$ , and therefore  $\Delta G_{\text{vdW}}$ , can be estimated from a knowledge of water density fluctuations in the absence of attractions.

Furthermore, recognizing that the presence of attractions in the core of a solute has no bearing on its hydration free energy, also allows for efficient estimation of hydration free energies in interfacial environments. When hydrating a solute in such inhomogeneous environments, attractions from neighboring solutes or interfaces acting on  $v$ , make it harder to empty the solute core. We recommend that such attractions should be turned off (at least partially) in a first step, with the corresponding free energy readily estimated using linear response theory. The solute core can then be emptied more easily in a second step because the free energy for doing so is substantially reduced.

## II. METHODS

### A. Models

To obtain a qualitative understanding of the influence of attractive potentials on density fluctuations, we focus on the solvation of a square well particle in the extended simple point charge (SPC/E) model of water [58]. While we focus on square-well solutes, our findings are general and hold for more realistic Lennard-Jones (LJ) solutes (see Supplementary Material [59]). We chose the SPC/E model for water because it reasonably mimics the bulk density, isothermal compressibility, surface tension, and liquid-vapor coexistence properties of real water; these properties primarily influence the behavior of water density fluctuations [18].

*Bulk Hydration:* The solutes we study interact with water oxygens through the pair potential shown in Fig-

ure 1a. The potential has a hard core of radius  $R_C$  and an attractive region of width  $R_S - R_C$  and depth  $\epsilon$ . We consider three core sizes spanning the small to large solute size regimes of hydrophobic solvation [34],  $R_C = 0.336$  nm,  $R_C = 0.6$  nm, and  $R_C = 0.9$  nm. The smallest solute corresponds to the effective hard sphere radius [23, 60] of a united atom methane with Lennard-Jones parameters,  $\sigma_{\text{LJ}} = 0.3448$  nm and  $\epsilon_{\text{LJ}} = 0.8956$  kJ/mol [61]. For all solutes, we choose a spherical shell region 0.3 nm in width, that is,  $R_S = R_C + 0.3$  nm, and set the well-depth,  $\beta\epsilon = 1$ .

While the repulsive part of the square-well potential is unambiguously given by the hard sphere pair potential,  $u_0$ , the attractive component can be defined in one of two ways (Figure 1). The first solute-water attractive potential,  $u_1(r)$ , arises from a WCA-like separation of a standard square-well potential [55] shown in Figure 1b, and acts on both the core and shell regions. The contribution of such solute-water attractions to the Hamiltonian of the system is

$$U_1(\bar{\mathbf{R}}) = -\epsilon(N_v(\bar{\mathbf{R}}) + N_{v_{\text{sh}}}(\bar{\mathbf{R}})), \quad (3)$$

and depends only on the number of molecules in the core and shell regions,  $N_v(\bar{\mathbf{R}})$  and  $N_{v_{\text{sh}}}(\bar{\mathbf{R}})$  respectively, where  $\bar{\mathbf{R}}$  denotes a configuration vector containing the positions of all the water oxygens. The results obtained from this potential can be qualitatively compared to the WCA attractive portion of the LJ potential; such a comparison is shown in the Supplementary Material [59]. Alternatively, an attractive potential that acts on the shell alone,  $u_2(r)$ , can be considered (Figure 1c), and its contribution to the system Hamiltonian is

$$U_2(\bar{\mathbf{R}}) = -\epsilon N_{v_{\text{sh}}}(\bar{\mathbf{R}}). \quad (4)$$

To vary the strength of attractions, we employ a linear coupling parameter in both cases, that is, we apply attractive potentials,  $\lambda_i U_i$ , and vary  $\lambda_i$  to change the interaction strength.

*Hydration at Interfaces:* We also consider the hydration of a hard solute in the vicinity of a square-well attractive surface. The surface excludes waters from a cuboid-shaped volume,  $v_s = 2.5 \times 2.5 \times 1.0$  nm<sup>3</sup>, and has attractive interactions of strength  $\lambda_1 \epsilon$  with water in an adjoining volume,  $v_1 = 2.5 \times 2.5 \times 0.3$  nm<sup>3</sup>. The cuboid-shaped hard solute,  $v = 1.5 \times 1.5 \times 0.3$  nm<sup>3</sup>, is then hydrated at the center of  $v_1$ .

## B. Simulation Details and Methods

All simulation data presented here were obtained using version 4.5.3 of the GROMACS molecular dynamics (MD) simulation package [62], modified to include the various external and biasing potentials used in this work. MD simulations were performed in the isothermal-isobaric (NPT) ensemble, where the canonical velocity-rescaling thermostat of Bussi *et al.* [63] was used to main-

tain a constant temperature of 300 K and a Parinello-Rahman barostat [64] was used to maintain a pressure of 1 bar. Electrostatic interactions were handled with the particle mesh Ewald method [65] with a real space cutoff of 1 nm and grid spacing of 0.12 nm. Similarly, van der Waals interactions were cutoff at a distance of 1 nm, and standard energy and pressure corrections were used to account for the influence of the truncated interactions [66].

*Bulk Hydration:* We denote the free energy of turning on the attractive field,  $\lambda_i U_i$  ( $i = 1, 2$ ), by  $\Delta G_i$ , and the fluctuations in the presence of the field by  $P_v^{(i)}(N)$ . For the specific case of square-well type potentials, the fields,  $U_1$  and  $U_2$ , depend only on  $N_v$  and  $N_{v_{\text{sh}}}$ , the number of water molecules in the solute core and hydration shell, respectively; that is  $U_i(\bar{\mathbf{R}}) = U_i(N_v(\bar{\mathbf{R}}), N_{v_{\text{sh}}}(\bar{\mathbf{R}}))$ . Thus, both  $\Delta G_i$  and  $P_v^{(i)}(N)$  can be readily obtained for a range of attractive strengths,  $\lambda_i$ , if the probability,  $P_{v, v_{\text{sh}}}(N, N_{\text{sh}})$ , of observing  $N$  and  $N_{\text{sh}}$  waters in  $v$  and  $v_{\text{sh}}$ , respectively, in the absence of any solute-water attractions is known. In particular,  $P_v^{(i)}(N)$  can be obtained from the exact expression

$$P_v^{(i)}(N) = C \sum_{N_{\text{sh}}=0}^{\infty} P_{v, v_{\text{sh}}}(N, N_{\text{sh}}) e^{-\beta \lambda_i U_i(N, N_{\text{sh}})}, \quad (5)$$

where  $C$  is a normalization constant. Similarly,  $\Delta G_i$  can be determined exactly from

$$\beta \Delta G_i = -\ln \sum_N \sum_{N_{\text{sh}}} P_{v, v_{\text{sh}}}(N, N_{\text{sh}}) e^{-\beta \lambda_i U_i(N, N_{\text{sh}})}. \quad (6)$$

Derivations of both Equation 5 and Equation 6 are included in the Supplementary Material [59]. To obtain  $P_{v, v_{\text{sh}}}(N, N_{\text{sh}})$ , we employed indirect umbrella sampling (INDUS) [27] with harmonic biasing potentials of the form

$$U_{\text{bias}}(\bar{\mathbf{R}}) = \frac{\kappa_1}{2} (\tilde{N}_v(\bar{\mathbf{R}}) - \tilde{N}_v^*)^2 + \frac{\kappa_2}{2} (\tilde{N}_{v_{\text{sh}}}(\bar{\mathbf{R}}) - \tilde{N}_{v_{\text{sh}}}^*)^2. \quad (7)$$

The INDUS approach dictates coarse-graining the discrete variables  $N_v$  and  $N_{v_{\text{sh}}}$  to the continuous variables  $\tilde{N}_v$  and  $\tilde{N}_{v_{\text{sh}}}$  through convolution with a truncated and shifted Gaussian; here we use a width of 0.01 nm and a cutoff of 0.03 nm for this smoothing function. The force constants,  $\kappa_i$ , and the biased potential minima,  $\tilde{N}_i^*$ , were tuned to achieve sufficient overlap between neighboring windows. The corresponding  $P_{v, v_{\text{sh}}}(N, N_{\text{sh}})$  distributions were then obtained from the biased simulations using the unbinned weighted histogram analysis method (UWHAM) [67, 68].

*Hydration at Interfaces:* Because simulating the square-well surface using MD would result in impulsive forces, we instead employ closely related continuous potentials to mimic the square-well potential, and make use of standard reweighting techniques [69] to estimate the behavior of a true square-well surface. The exclusion of water molecules from the square-well surface is accomplished by applying the potential  $\tilde{U}_s(\bar{\mathbf{R}}) = \phi_s \tilde{N}_{v_s}(\bar{\mathbf{R}})$ ,

where  $\tilde{N}_{v_s}(\bar{\mathbf{R}})$  is the coarse-grained number of waters in  $v_s$ , and  $\beta\phi_s = 12$  was chosen to yield a substantial number of configurations with the volume empty. Analogously, the attractive potential in the adjoining hydration shell volume,  $v_1$ , is  $\tilde{U}_1(\bar{\mathbf{R}}) = \phi_1 \tilde{N}_{v_1}(\bar{\mathbf{R}})$ , where  $\tilde{N}_{v_1}$  is the coarse-grained number of water molecules in  $v_1$ , and  $\phi_1 \equiv -\lambda_1 \epsilon$ . Both  $\tilde{N}_{v_s}$  and  $\tilde{N}_{v_1}$ , were defined according to the INDUS prescription described above. To estimate the probability,  $P_v^{(1)}(N)$ , of observing  $N$  water molecules in  $v$ , a series of INDUS simulations were then performed using a biasing potential similar to the one in Equation 7, but with  $\kappa_2 = 0$ . Averages in a system with a true square-well surface,  $\langle(\dots)\rangle$ , can be related to averages in the mimic square-well system,  $\langle(\dots)\rangle'$  (simulated using the potentials  $U_s$  and  $U_1$ ), through reweighting as,

$$\langle(\dots)\rangle = \frac{\left\langle (\dots) \delta_{0, N_{v_s}(\bar{\mathbf{R}})} e^{\beta\phi_s \tilde{N}_{v_s}(\bar{\mathbf{R}}) - \beta\phi_1 [N_{v_1}(\bar{\mathbf{R}}) - \tilde{N}_{v_1}(\bar{\mathbf{R}})]} \right\rangle'}{\left\langle \delta_{0, N_{v_s}(\bar{\mathbf{R}})} e^{\beta\phi_s \tilde{N}_{v_s}(\bar{\mathbf{R}}) - \beta\phi_1 [N_{v_1}(\bar{\mathbf{R}}) - \tilde{N}_{v_1}(\bar{\mathbf{R}})]} \right\rangle'}, \quad (8)$$

where  $\delta_{i,j}$  is the Kronecker delta function, and  $N_{v_s}$  and  $N_{v_1}$  are the number of waters in  $v_s$  and  $v_1$  respectively. This average, in turn, is evaluated using UWHAM [67, 68]. The probability,  $P_v^{(2)}(N)$ , of finding  $N$  waters in  $v$ , in the absence of favorable surface-water interactions in  $v$ , is obtained through another reweighting,

$$P_v^{(2)}(N) = C' P_v^{(1)}(N) e^{-\beta\lambda_1 \epsilon N}, \quad (9)$$

where  $C'$  is a normalization constant.

### III. THE INFLUENCE OF ATTRACTIONS ON WATER DENSITY FLUCTUATIONS

The attractive field,  $\lambda_1 U_1$ , which acts on both the core and shell regions, significantly affects water density fluctuations in the core, as shown in Figure 2a for  $R_C = 0.6$  nm. As expected, the average number of molecules in the core region,  $\langle N_v(\bar{\mathbf{R}}) \rangle_{\lambda_1}$ , increases with increasing field strength  $\lambda_1$ . Here,  $\langle(\dots)\rangle_{\lambda_i}$  represents an ensemble average in the presence of the attractive field  $\lambda_i U_i$ . Importantly, the non-Gaussian nature, or the fatness of the low- $N$  tails is also diminished, making it increasingly difficult to empty the core region as  $\lambda_1$  is increased. To better understand how water density fluctuations are affected by  $\lambda_1 U_1$ , we extend the model for density fluctuations proposed by Huang and Chandler [42] (See Appendix A). The Huang-Chandler model assumes that reducing  $N_v$  below its average value results in the formation of a bubble, and that the low- $N$  tail in the water density fluctuations can be described by the energetics of growing the bubble. As shown in the Appendix A, in the presence of  $\lambda_1 U_1$ , additional work has to be done against the attractive field to expand the bubble, which results in the low- $N$  tails of the  $P_v^{(1)}(N)$  distributions being diminished. This qualitative behavior is also

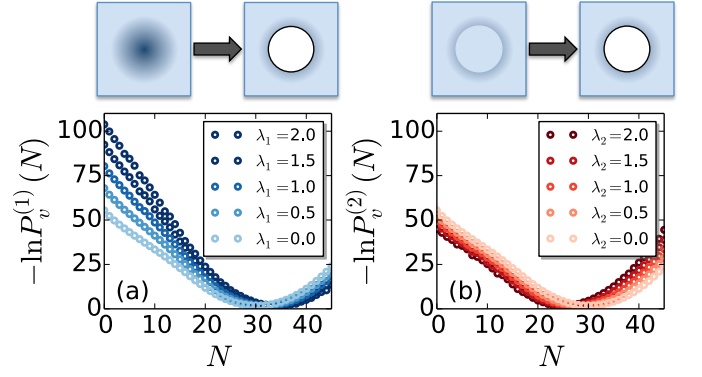


FIG. 2: (a) The presence of an attractive field that acts on *both* the core and shell regions,  $\lambda_1 U_1$ , drastically alters density fluctuations in the core region, as shown here for a spherical solute with  $R_C = 0.6$  nm and  $R_S = 0.9$  nm. On increasing the attractive strength  $\lambda_1$ , the tails at low  $N$  become steeper, indicating that it is more difficult to evacuate the core volume. (b) In contrast, fluctuations in the core region are not significantly perturbed by the attractive field,  $\lambda_2 U_2$ , which acts on the hydration shell alone. As the strength of the attractive interactions,  $\lambda_2$ , is increased, only a small *decrease* in the average number of molecules in the core is observed.

observed for both smaller and larger solutes (See Supplementary Material [59]).

In contrast, in the presence of an attractive field such as  $\lambda_2 U_2$ , which acts only in the hydration shell, the nature of the density fluctuations in the core region is not changed; see Figure 2b. Only a small decrease in the mean,  $\langle N_v(\bar{\mathbf{R}}) \rangle_{\lambda_2}$ , is observed when the attractive strength,  $\lambda_2$ , is increased. In Figure 3, the fluctuation spectra,  $P_v^{(2)}$ , are shifted horizontally so that their means are aligned. For all the  $R_C$ -values studied, the shifted spectra,  $P_v^{(2)}(\mathcal{N}) = P_v^{(2)}(N - \Delta N)$ , where  $\Delta N \equiv \langle N_v(\bar{\mathbf{R}}) \rangle_{\lambda_2} - \langle N_v(\bar{\mathbf{R}}) \rangle_0$ , corresponding to various attractions,  $\lambda_2$ , collapse onto a universal curve. Remarkably, this invariance of water density fluctuations around the mean is independent of whether the fluctuations are nearly Gaussian (small  $R_C$ ) or have pronounced fat tails (large  $R_C$ ). This is the central result of this work.

To understand the basis of this invariance, we recognize that the addition of a linear potential (such as  $\lambda_2 N_{v_{sh}}$ ) to a perfectly harmonic basin (such as one arising from  $P_{v,v_{sh}}$  being Gaussian) simply translates the basin (and hence the entire distribution) horizontally [48]. Attractions in the shell favor configurations with more waters than the average, coupling to the high- $N_{sh}$  region of  $P_{v,v_{sh}}(N, N_{sh})$ . Such high- $N$  regions of water number distributions in bulk water have previously been shown to follow Gaussian statistics [26, 27], and  $P_{v,v_{sh}}(N, N_{sh})$  similarly follows a bivariate normal distribution with nonzero correlation between  $N$  and  $N_{sh}$  near its mean.

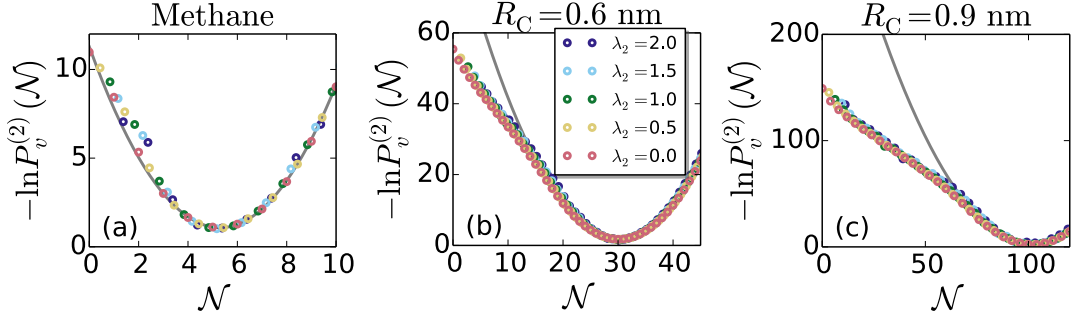


FIG. 3: Water number distributions,  $P_v^{(2)}$ , in the presence of attractive fields  $\lambda_2 U_2$ , shifted horizontally to align their means,  $\langle N_v(\mathbf{R}) \rangle_{\lambda_2}$ , are shown for probe volumes with radii (a) equivalent to that of a methane, (b) 0.6 nm, and (c) 0.9 nm. Fluctuation spectra obtained over a range of attractions collapse onto a single curve when plotted as a function of  $\mathcal{N} \equiv N - \Delta N$ , where  $\Delta N \equiv \langle N_v(\mathbf{R}) \rangle_{\lambda_2} - \langle N_v(\mathbf{R}) \rangle_0$ , highlighting that the presence of an attractive potential in the hydration shell simply decreases the average number of water molecules in the core, and not the fluctuations around the mean. This is true independent of solute size and whether the fluctuations follow Gaussian statistics (solid gray lines).

Thus, the  $\lambda_2 U_2$  potential, which is linear in  $N_{v_{\text{sh}}}(\mathbf{R})$ , and couples to the Gaussian high- $N_{\text{sh}}$  region of  $P_{v,v_{\text{sh}}}(N, N_{\text{sh}})$  is expected to shift the entire  $P_{v,v_{\text{sh}}}(N, N_{\text{sh}})$  distribution towards higher  $N_{\text{sh}}$ -values. As  $\lambda_2$  is increased, the correlations between  $N_{v_{\text{sh}}}(\mathbf{R})$  and  $N_v(\mathbf{R})$  then alter  $\langle N_v(\mathbf{R}) \rangle$ , but leave the remainder of  $P_v^{(2)}(N)$  unaltered.

While the application of  $U_2$  increases water density in the shell region, packing effects can lead to a concomitant *decrease* in the average number of waters in the core region. The layering of water density at the core-shell interface, shown in Figure 4a, highlights that attractions in the shell indeed decrease the average density in the core. Interestingly, as shown in Figure 4b, this decrease in the average number of molecules in the core region upon application of the potential  $U_2$  can be accurately estimated using linear response theory,

$$\Delta N \approx -\lambda_2 \beta \epsilon \langle \delta N_v(\mathbf{R}) \delta N_{v_{\text{sh}}}(\mathbf{R}) \rangle_0, \quad (10)$$

given the correlation,  $\langle \delta N_v(\mathbf{R}) \delta N_{v_{\text{sh}}}(\mathbf{R}) \rangle_0$ , between fluctuations in  $N_v(\mathbf{R})$  and  $N_{v_{\text{sh}}}(\mathbf{R})$  in the absence of attractions. As discussed in section V, this agreement will facilitate the development of approximations for estimating hydration free energies of realistic solutes in bulk water. For now, having quantified the ease with which the core can be emptied in the presence of the attractive fields  $\lambda_1 U_1$  and  $\lambda_2 U_2$ , we turn to the free energies of turning on these potentials in bulk water.

#### IV. ATTRACTIONS IN THE SOLUTE CORE DO NOT INFLUENCE ITS HYDRATION FREE ENERGY

Two alternative paths for hydrating the square-well solute of Figure 1a are shown in Figure 5a. In the upper path, the attractive potential  $\lambda_1 U_1$  is first turned on in

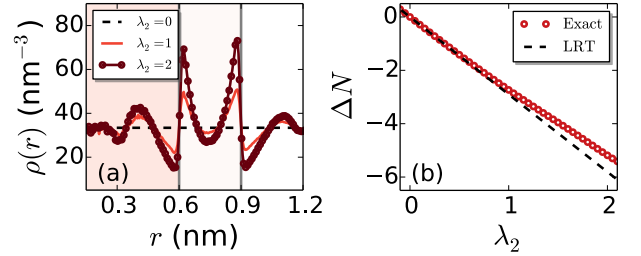


FIG. 4: (a) Attractions in the shell region, indicated by vertical gray lines, increase water density in the shell, but result in a decrease in the average water density in the core region, as shown here for  $R_C = 0.6$  nm. The decrease in the core density on increasing  $\lambda_2$  is a direct result of layering at the interface between the core and shell regions. (b) This decrease in the average number of water molecules in the core region,  $\Delta N$ , as the strength of the potential  $\lambda_2$  is increased, can be accurately predicted with linear response theory (LRT), Equation 10.

the core and the shell, whereas in the lower path, the potential,  $\lambda_2 U_2$ , is turned on in the shell alone. We denote the free energy for turning on the attractive field,  $\lambda_i U_i$ , in bulk water by  $\Delta G_i$  ( $i = 1, 2$ ). Then, in the second step, a cavity is created in the core region in the presence of the corresponding attractive potential. The free energy for creating such a cavity is given by  $-k_B T \ln P_v^{(i)}(N = 0)$ , and is informed by the statistics of density fluctuations,  $P_v^{(i)}(N)$ , shown in Figure 2.

As shown in Figures 5b and c, turning on attractions in the both the core and shell regions results in a large and negative free energetic gain  $\Delta G_1$ , whereas the corresponding free energy,  $\Delta G_2$ , for turning on attractions in the shell alone, is significantly smaller in magnitude. However, the total hydration free energy is



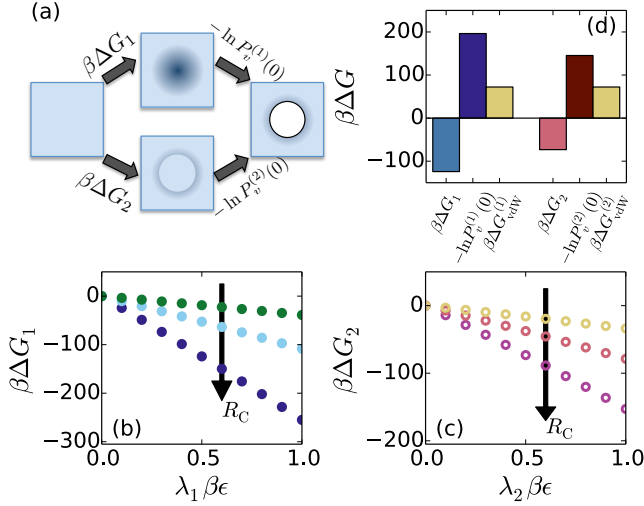


FIG. 5: (a) The hydration of a realistic hydrophobic solute is carried out in two steps: an attractive field,  $\lambda_i U_i$ , corresponding to the solute-water attractions is first turned on in bulk water, and water is emptied from the solute core in a subsequent step. The corresponding free energies are denoted by  $\beta\Delta G_i$  and  $-\ln P_v^{(i)}(0)$  respectively, where  $i = 1$  represents attractions in both the core and shell regions (top) and  $i = 2$  represents attractions in only the shell region (bottom). (b, c) The free energy of turning on the attractive potential  $U_1$  is much larger than that for turning on  $U_2$ . (d) This results in a larger cancellation between the favorable  $\beta\Delta G_i$  and the unfavorable  $-\ln P_v^{(i)}(0)$  free energies for  $i = 1$ , because the total hydration free energy  $\Delta G_{\text{vdW}} = \Delta G_i - \ln P_v^{(i)}(0)$  is the same for both paths, as shown here for  $R_C = 0.9$  nm and  $\lambda_i\beta\epsilon = 0.5$ .

given by the sum of the favorable  $\Delta G_i$  and unfavorable  $-k_B T \ln P_v^{(i)}(N=0)$  terms. The diminished fluctuations in the presence of  $U_1$  (Figure 2a) mean that the unfavorable free energy to form a cavity is also larger in the presence of  $U_1$ . The total hydration free energy,  $\Delta G_{\text{vdW}}$ , estimated through the upper path thus results in a significant cancellation between the favorable  $\Delta G_1$  and the unfavorable  $-k_B T \ln P_v^{(1)}(N=0)$  terms, as shown in Figure 5d. Such large cancellations can lead to significant numerical uncertainty in the estimation of  $\Delta G_{\text{vdW}}$ .

In contrast, the lower path in Figure 5a reduces the extent of such a cancellation by recognizing that the presence or absence of attractions in the solute core plays no role in determining the overall solute hydration free energy, and minimizing the volume over which the attractions act. In principle, the cancellation between the favorable  $\beta\Delta G_i$  and unfavorable  $-\ln P_v^{(i)}(N=0)$  terms can be further reduced or even eliminated by turning on repulsions in the core in conjunction with attractions in the shell in the first step. Even more complex, low-variance pathways to hydration may also be chosen to cleverly minimize the computational overhead [70, 71]. However, the advantage of the paths considered here is

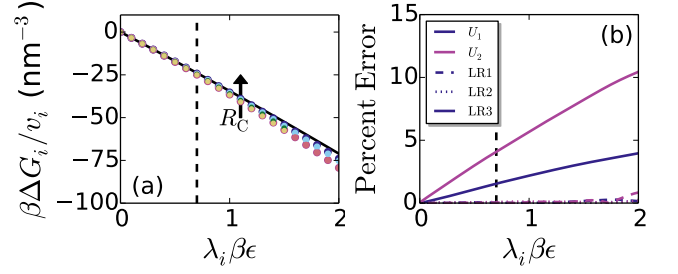


FIG. 6: (a) The free energy,  $\Delta G_i$ , of turning on the attractive potential, normalized by the volume,  $v_i$ , on which the attractions act, is largely independent of solute size and the range over which the potential acts.  $\Delta G_i$  is well described by the linear response prediction LR3 of Equation 13 (solid black line). (b) The three linear response theory estimates for  $\Delta G_i$  are found to be highly accurate, as shown here for the  $R_C = 0.9$  nm solute. The dashed line indicates a typical magnitude of attractions for realistic solutes [72].

that they allow for particularly simple analytical approximations for  $\Delta G_i$ , as discussed in the next section.

## V. EFFICIENT ESTIMATION OF BULK HYDRATION FREE ENERGIES

Because attractive fields couple to the Gaussian high- $N$  tails of water density distributions, we expect linear response theory to provide an accurate estimate of the free energy for turning on such fields. In particular, we consider the following approximate forms of Equation 6:

$$\beta\Delta G_i \approx \frac{\beta\lambda_i}{2} \left[ \langle U_i(\mathbf{R}) \rangle_0 + \langle U_i(\mathbf{R}) \rangle_{\lambda_i} \right] \quad (11)$$

$$\approx \beta\lambda_i \langle U_i(\mathbf{R}) \rangle_0 - \frac{\beta^2 \lambda_i^2}{2} \langle (\delta U_i(\mathbf{R}))^2 \rangle_0, \quad (12)$$

referred to as LR1 and LR2, respectively. In the context of thermodynamic integration, LR1 represents the application of a trapezoid rule to integrate over the entire range of  $\lambda_i$ , while LR2 uses the initial slope of the thermodynamic force at  $\lambda_i = 0$  and extrapolates over the range of  $\lambda_i$  to integrate. For the square-well potentials considered here, the free energy of turning on the attractive external field can further be approximated analytically, yielding

$$\frac{\beta\Delta G_i}{v_i} \approx -\rho_B(\lambda_i\beta\epsilon) - \frac{\rho_B^2 \kappa_T k_B T}{2} (\lambda_i\beta\epsilon)^2, \quad (13)$$

where  $v_i$  is the volume over which the potential is applied, and  $\rho_B$  and  $\kappa_T$  are the density and isothermal compressibility of bulk water, respectively. To derive Equation 13 (referred to as LR3) from LR2, we recognize that fluctuations in  $U_i$  are related to fluctuations in water density, which in turn are related to the isothermal compressibility. The exact derivations of LR3 and its generic form

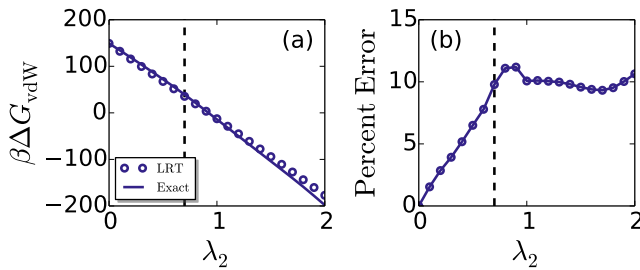


FIG. 7: (a) The total hydration free energy of a realistic hydrophobic solute,  $\Delta G_{\text{vdW}} = \Delta G_2 - k_B T \ln P_v^{(2)}(N=0)$ , can be described with good accuracy using the approximations in Equations 13 and 14, as shown here for the solute with  $R_C = 0.9$  nm. (b) Indeed, for realistic values [72] of  $\lambda_2$  (dashed line), the error in  $\Delta G_{\text{vdW}}$  is less than ten percent.

applicable to slowly-varying potentials are provided in the Supplementary Material [59].

LR3 suggests that  $\Delta G_i$  can be estimated simply from a knowledge of the strength of the attractions, the volume over which they act, and bulk properties of the solvent. It predicts that  $\Delta G_i$  is proportional to the volume,  $v_i$ , over which the attractive interactions act, and is quadratic in the strength of the attractions,  $\lambda_i$ , consistent with previous findings [73]. We find LR3 to be true to a very good approximation, as illustrated in Figure 6; it begins to break down only for small volumes and large  $\lambda_i$  (see Supplementary Material [59]). The discrepancy between the predicted and simulation results for typical attraction strengths [72] is less than five percent. The approximations, LR1 and LR2, provide even more accurate estimates of  $\Delta G_i$ , with errors of less than a percent.

In addition to being able to obtain accurate estimates for  $\Delta G_i$  from linear response theory, our essential finding that attractions (in the hydration shell of a solute) do not affect water density fluctuations, also permits accurate estimates of the free energy for emptying the solute core,  $-\ln P_v^{(2)}(N=0)$ . Turning on attractive interactions in the hydration shell only results in a small shift in the mean,  $\Delta N$  (Figure 3), which itself can be readily estimated using linear response theory (Equation 10), so that fluctuations in the presence and absence of attractions can be related through

$$P_v^{(2)}(N) \approx P_v(N + \Delta N). \quad (14)$$

Evaluating the components of the total bulk hydration free energy (Equation 2) using Equation 13 and Equation 14 (for  $N=0$ ) leads to accurate predictions for  $\Delta G_{\text{vdW}}$ , as shown in Figure 7 for the  $R_C = 0.9$  nm solute. This approach thus presents an efficient way to accurately estimate bulk hydration free energies of realistic attractive solutes,  $\Delta G_{\text{vdW}}$ , starting from information on the statistics of water density fluctuations in bulk water,  $P_v(N)$ .

## VI. EFFICIENT ESTIMATION OF HYDRATION FREE ENERGIES IN HETEROGENEOUS ENVIRONMENTS

Figure 5d highlights that it is efficient to turn off any attractive fields in the solute core before estimating the free energy for emptying the core. Such attractive fields (acting on water) can originate not only from the solute itself, but also from neighboring solutes or interfaces. To demonstrate this, in Figure 8a, we illustrate the hydration of a hard solute (shown in green) in the vicinity of a surface (shown in purple) that interacts with water through an attractive square well potential (dark blue gradient). The attractive interactions from the square well act on the core region of the solute, making it harder to empty (upper path). Alternatively, the attractive potential can first be turned off in the solute core region, as shown in the lower path of Figure 8a, with the corresponding free energy accurately estimated by linear response theory. Turning off the attractions also ensures that emptying the solute cavity is easier, making it the more efficient alternative (Figure 8c).

Indeed, the free energy needed to create a cavity in the presence of the attractive field is significantly higher (roughly  $30 k_B T$  versus  $10 k_B T$ ), as seen in Figure 8b. When the attractive field is turned off in the solute core, the average number of water molecules in the core decreases and a significant non-Gaussian tail emerges in  $P_v^{(2)}(N)$  at low  $N$ ; both effects act in concert to lower the free energy needed to form a cavity. Such a lowering of the cavity formation free energy allows for more efficient estimation of hydration free energies in heterogeneous environments, because fewer biased simulations are needed to probe the entire range of density fluctuations in the solute core.

*Caveat:* While the free energy of turning off the attractive interactions inside the solute core was accurately described using linear response theory (LR1) here, this may not always be the case. In particular, turning off attractions completely in a sufficiently large volume adjacent to an extended hydrophobic surface can dewet the volume, causing linear response to break down [56]. However, only weak attractions are needed to wet a hydrophobic surface; recent work suggests that an attractive strength of roughly  $0.1 k_B T$  [74, 75] is sufficient to prevent dewetting. Thus, attractions may be turned off partially without triggering dewetting, and the corresponding free energy safely estimated using linear response. In such cases, our revised recommendation is to reduce the surface-water attractions down to  $0.1 k_B T$  in a first step, followed by cavity creation in the presence of reduced attractions.

## VII. CONCLUSIONS

Water number fluctuations in small volumes follow Gaussian statistics, while those in large volumes have non-Gaussian tails at low density. An understanding of

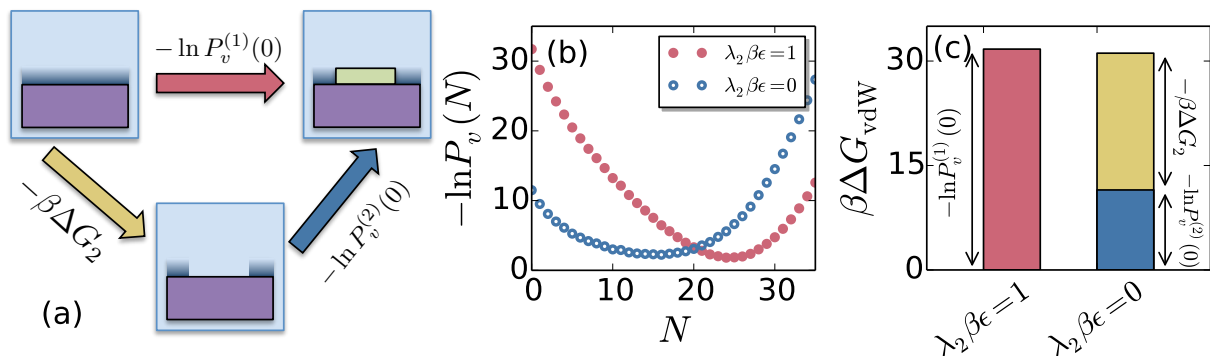


FIG. 8: (a) Top: The hydration of a  $1.5 \times 1.5 \times 0.3 \text{ nm}^3$  cuboid-shaped hard solute (green) adjacent to a  $2.5 \times 2.5 \times 1.0 \text{ nm}^3$  attractive square-well surface (purple, well depth  $\beta\epsilon = 1$  and width  $0.3 \text{ nm}$ ), entails the creation of a cavity in the presence of attractive surface-water interactions. Bottom: Alternatively, the solute can be hydrated in two steps; turning off the attractions in the core of the solute in the first step, followed by creating a cavity in the second step. (b) Emptying the solute core requires a significantly smaller free energy in the absence of attractions in the core. (c) While the solute hydration free energy is the same in both cases, the free energy for turning off the attractions,  $-\beta\Delta G_2$ , can be accurately estimated by using linear response theory, and the free energy for subsequently emptying the solute core,  $-\ln P_v^{(2)}(0)$ , is smaller than  $-\ln P_v^{(1)}(0)$ , and therefore easier to estimate.

these fluctuations has made seminal contributions to our understanding of the hydration and interactions of idealized, purely repulsive hydrophobic solutes. To similarly inform the hydration of realistic hydrophobic solutes that have attractive interactions with water, here we have provided a description of these fluctuations in the presence of attractive fields. We find that WCA-like attractive potentials [55], which are non-zero both inside and outside the solute core, alter the nature of water density fluctuations significantly, making it more difficult to create a cavity. In contrast, when attractions are turned on in the hydration shell alone, water density fluctuations are largely unaltered in both small and large volumes.

We also find that the favorable free energy for turning on the various attractive fields is readily approximated by linear response theory and is proportional to the volume on which the fields act. Consequently, the free energy of turning on attractions in both the core and the shell is much larger than the free energy of turning on attractions in the hydration shell alone. When attractions are turned on in both the core and the shell, the favorable free energy of turning on those attractions and the unfavorable free energetics of emptying the solute core in the presence of those attractions, largely cancel in the estimation of the total solute hydration free energy. This cancelation is minimized when the solute-solvent attractions are turned on in the hydration shell alone, making it the more efficient route for estimating hydration free energies of realistic solutes.

In addition to informing hydrophobic hydration and interactions, an understanding of density fluctuations has also facilitated the development of novel simulation methods. By leveraging an understanding of water density fluctuations, and the associated response of water density to perturbations, Patel and Garde recently introduced a method for estimating cavity hydration free energies that is two orders of magnitude more efficient as com-

pared with umbrella sampling [29]. Our characterization of density fluctuations in the presence of slowly-varying attractive interactions has similarly allowed us to suggest strategies for more efficiently estimating hydration free energies of realistic solutes, both in bulk water and in the vicinity of interfaces.

Our approach involves turning on attractions in bulk water, followed by emptying the solute core, and is in contrast with traditional methods that first create a cavity in bulk water, and subsequently turn on attractions. Because attractions have little effect on the structure of water adjacent to small cavities (small enough for the hydrogen bond network of water to be maintained around them) [23, 28, 73, 76], traditional methods can readily estimate the free energy for turning on attractions using linear response theory. However, larger cavities induce dewetting in their vicinity; attractions rewet the solute, so that water density and consequently the solute-water interaction energy do not vary linearly with the strength of attractions [1]. Indeed, recent work from Underwood and Ben-Amotz [56] has shown a transition from linear to non-linear response as the solute size is increased [56], occurring at roughly  $1 \text{ nm}$ ; the length scale corresponding to the crossover in hydrophobic hydration [34]. Because our approach involves turning on attractions in bulk water *before* creating a solute core, it circumvents the breakdown of linear response theory that plagues the traditional perturbative treatment of attractive interactions.

The approach to solvation presented here is similar in spirit to the two-step method considered by Weeks and coworkers [35, 77] in their development of the molecular-scale van der Waals (MVDW) theory for inhomogeneous systems, which was subsequently combined with the Gaussian field model of Chandler [31] to yield the LCW theory of hydrophobicity [34]. The MVDW theory first considers turning on all slowly-varying interactions in the bulk fluid, which include not only the solute-solvent



attractions that we consider here, but also long-ranged *solvent-solvent* interactions. Near large hydrophobic solutes, the long-ranged solvent-solvent attractive interactions are unbalanced and result in a net repulsion away from the solute [35]. Such repulsions couple to the non-Gaussian low  $N$  tails in water number distributions and result in dewetting; thus their effect can not be captured using linear response. Recognizing this, the LCW [34] theory employed a Landau-Ginzburg free energy functional that enables repulsive potentials to induce dewetting. In contrast, because our focus here has been on slowly-varying attractive potentials that couple to the Gaussian high- $N$  region of water number distributions, we are able to make judicious use of linear response theory.

The results presented here rely on two essential properties of attractive interactions: (i) they are slowly varying and thereby minimally perturb the structure of water, and (ii) they couple to Gaussian tails of water number distributions. Thus, the lessons learned from this work will also be applicable to other interactions that possess these two characteristic features. In particular, recent work by Weeks and co-workers [28, 78–82] has shown that it is instructive to resolve electrostatic interactions into short- and long-ranged components, in analogy with the WCA prescription, which divides the Lennard-Jones potential into a short-ranged repulsive and a slowly-varying attractive component [55]. While the long-ranged electrostatic component follows linear response theory, underpinning dielectric continuum theories [83], the short-ranged interactions are more complex, leading to directional hydrogen bonds and specific ion effects [84, 85]. Thus, in hydrating ions, for example, it would be instructive to turn on the long-ranged component of electrostatic interactions first, and investigate how it affects water density fluctuations; such calculations will be the focus of future work.

## ACKNOWLEDGMENTS

The authors acknowledge partial financial support from the National Science Foundation through a Seed Grant from the University of Pennsylvania Materials Research Science and Engineering Center (NSF UPENN MRSEC DMR 11-20901).

## Appendix A: Extension of the Huang-Chandler Model

Here, we describe the response of  $P_v^{(1)}(N)$  to an attractive potential by extending the model developed by Huang and Chandler (HC) [42]. The model assumes that reducing  $N_v(\bar{\mathbf{R}})$  below its average value results in the formation of a spherical bubble of radius  $r_b$ , and approximates the density outside the bubble to be the bulk density,  $\rho_B$ . The free energetics of density fluctuations (in the absence of an external field) are then given by the work that must be performed to grow the bubble against the surface tension  $\gamma$  and the external pressure  $\Delta\mathcal{P} = \mathcal{P} - \mathcal{P}^{\text{sat}}$ , where  $\mathcal{P}$  and  $\mathcal{P}^{\text{sat}}$  are the system pressure and the saturation pressure respectively. Thus,

$$\begin{aligned} -k_B T \ln P_v(N) &= G_0(r_b(N)) \\ &\approx \frac{4\pi}{3} r_b^3(N) \Delta\mathcal{P} + 4\pi r_b^2(N) \gamma. \end{aligned} \quad (\text{A1})$$

For the sake of simplicity, the term corresponding to the translational entropy of the bubble has been omitted from the above expression. The radius of the bubble  $r_b$  is related to  $N$  by

$$r_b(N) = \left[ \frac{3}{4\pi\rho_B} (\langle N_v(\bar{\mathbf{R}}) \rangle_0 - N) \right]^{1/3}. \quad (\text{A2})$$

In the presence of attractions in  $v$ , work must also be done against the attractive potential, such that

$$-\ln P_v^{(1)}(N) = -\ln P_v(N) + \frac{4\pi}{3} r_b^3(N) \rho_B \lambda_1 \beta \epsilon \quad (\text{A3})$$

$$= \frac{4\pi}{3} r_b^3(N) \beta \Delta\tilde{\mathcal{P}} + 4\pi r_b^2(N) \beta \gamma, \quad (\text{A4})$$

where we have defined

$$\Delta\tilde{\mathcal{P}} \equiv \Delta\mathcal{P} + \lambda_1 \epsilon \rho_B \quad (\text{A5})$$

as an effective pressure in the probe volume, to illustrate that an attractive potential that couples to  $N_v(\bar{\mathbf{R}})$  affects  $P_v(N)$  in the same manner as the pressure. Higher  $\Delta\tilde{\mathcal{P}}$  results in an increase in magnitude of the slope at low  $N$ , consistent with the results in Figure 2a. Conversely, repulsive potentials, corresponding to negative values of  $\lambda_1$ , should decrease  $\Delta\tilde{\mathcal{P}}$  and make it easier to create a cavity in the liquid.

- 
- [1] D. Chandler, *Nature* **437**, 640 (2005).
  - [2] P. Ball, *Chem. Rev.* **108**, 74 (2008).
  - [3] B. J. Berne, J. D. Weeks, and R. Zhou, *Ann. Rev. Phys. Chem.* **60**, 85 (2009).
  - [4] S. N. Jamadagni, R. Godawat, and S. Garde, *Ann. Rev. Chem. Biomol. Engg.* **2**, 147 (2011).
  - [5] N. T. Southall, K. A. Dill, and A. D. J. Haymet, *J. Phys. Chem. B* **106**, 521 (2002).
  - [6] Y. Levy and J. N. Onuchic, *Annu Rev Biophys Biomol Struct* **35**, 389 (2006).
  - [7] D. Thirumalai, E. P. O'Brien, G. Morrison, and C. Hyeon, *Annu Rev Biophys* **39**, 159 (2010).
  - [8] C. M. Dobson, *Nature* **426**, 884 (2003).
  - [9] M. G. Krone, L. Hua, P. Soto, R. Zhou, B. J. Berne, and J.-E. Shea, *J. Am. Chem. Soc.* **130**, 11066 (2008).
  - [10] D. Thirumalai, G. Reddy, and J. E. Straub, *Acc. Chem*

- Res. **45**, 83 (2012).
- [11] G. M. Whitesides and B. Grzybowski, *Science* **295**, 2418 (2002).
  - [12] E. Rabani, D. R. Reichman, P. L. Geissler, and L. E. Brus, *Nature* **426**, 271 (2003).
  - [13] J. A. Morrone, J. Li, and B. J. Berne, *J. Phys. Chem. B* **116**, 378 (2012).
  - [14] C. Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membranes* (John Wiley & Sons, 1973).
  - [15] L. Maibaum, A. R. Dinner, and D. Chandler, *J. Phys. Chem. B* **108**, 6778 (2004).
  - [16] G. Hummer, S. Garde, A. E. García, M. E. Paulaitis, and L. R. Pratt, *J. Phys. Chem. B* **102**, 10469 (1998).
  - [17] H. S. Ashbaugh and L. R. Pratt, *Rev. Mod. Phys.* **78**, 159 (2006).
  - [18] D. Chandler and P. Varilly, in *Proceeding of the International School of Physics "Enrico Fermi"*, Vol. 176, edited by F. Mallamace and H. E. Stanley (IOS, Amsterdam; SIF, Bologna, 2012) pp. 75–111.
  - [19] B. Widom, *Science* **157**, 375 (1967).
  - [20] F. H. Stillinger, *J. Solution Chem.* **2**, 141 (1973).
  - [21] G. E. Crooks and D. Chandler, *Phys. Rev. E* **56**, 4217 (1997).
  - [22] D. M. Huang, P. L. Geissler, and D. Chandler, *J. Phys. Chem. B* **105**, 6704 (2001).
  - [23] D. M. Huang and D. Chandler, *J. Phys. Chem. B* **106**, 2047 (2002).
  - [24] M. V. Athawale, G. Goel, T. Ghosh, T. M. Truskett, and S. Garde, *Proc. Nat. Acad. Sci.* **104**, 733 (2007).
  - [25] G. Goel, M. V. Athawale, S. Garde, and T. M. Truskett, *J. Phys. Chem. B* **112**, 13193 (2008).
  - [26] A. J. Patel, P. Varilly, and D. Chandler, *J. Phys. Chem. B* **114**, 1632 (2010).
  - [27] A. J. Patel, P. Varilly, D. Chandler, and S. Garde, *J. Stat. Phys.* **145**, 265 (2011).
  - [28] R. C. Remsing and J. D. Weeks, *J. Phys. Chem. B* **117**, 15479 (2013).
  - [29] A. J. Patel and S. Garde, *J. Phys. Chem. B* **118**, 1564 (2014).
  - [30] L. R. Pratt and D. Chandler, *J. Chem. Phys.* **67**, 3683 (1977).
  - [31] D. Chandler, *Phys. Rev. E* **48**, 2898 (1993).
  - [32] G. Hummer, S. Garde, A. E. Garcia, A. Pohorille, and L. R. Pratt, *Proc. Nat. Acad. Sci.* **93**, 8951 (1996).
  - [33] S. Garde, G. Hummer, A. E. Garcia, M. E. Paulaitis, and L. R. Pratt, *Phys. Rev. Lett.* **77**, 4966 (1996).
  - [34] K. Lum, D. Chandler, and J. D. Weeks, *J. Phys. Chem. B* **103**, 4570 (1999).
  - [35] J. D. Weeks, *Ann. Rev. Phys. Chem.* **53**, 533 (2002).
  - [36] P. Varilly, A. J. Patel, and D. Chandler, *J. Chem. Phys.* **134**, 074109 (2011).
  - [37] B. Widom, *J. Chem. Phys.* **39**, 2808 (1963).
  - [38] T. L. Beck, M. E. Paulaitis, and L. R. Pratt, *The Potential Distribution Theorem and Models of Molecular Solutions* (Cambridge University Press, 2006).
  - [39] P. R. ten Wolde, *J. Phys. Cond. Matt.* **14**, 9445 (2002).
  - [40] L. R. Pratt, *Ann. Rev. Phys. Chem.* **53**, 409 (2002).
  - [41] G. Hummer, S. Garde, A. Garcia, M. Paulaitis, and L. Pratt, *Proc. Natl. Acad. Sci. USA* **95**, 1552 (1998).
  - [42] D. M. Huang and D. Chandler, *Phys. Rev. E* **61**, 1501 (2000).
  - [43] J. Mittal and G. Hummer, *Proc. Nat. Acad. Sci.* **105**, 20130 (2008).
  - [44] S. Sarupria and S. Garde, *Phys. Rev. Lett.* **103**, 037803 (2009).
  - [45] R. Godawat, S. N. Jamadagni, and S. Garde, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15119 (2009).
  - [46] H. Acharya, S. Vembanur, S. N. Jamadagni, and S. Garde, *Faraday Discuss.* **146**, 353 (2010).
  - [47] A. J. Patel, P. Varilly, S. N. Jamadagni, H. Acharya, S. Garde, and D. Chandler, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17678 (2011).
  - [48] A. J. Patel, P. Varilly, S. N. Jamadagni, M. F. Hagan, D. Chandler, and S. Garde, *J. Phys. Chem. B* **116**, 2498 (2012).
  - [49] R. Zhou, X. Huang, C. J. Margulis, and B. J. Berne, *Science* **305**, 1605 (2004).
  - [50] P. Liu, X. Huang, R. Zhou, and B. J. Berne, *Nature* **437**, 159 (2005).
  - [51] N. Choudhury and B. M. Pettitt, *J. Am. Chem. Soc.* **129**, 4847 (2007).
  - [52] N. Choudhury and B. M. Pettitt, *Mol. Sim.* **31**, 457 (2005).
  - [53] J. C. Rasaiah, S. Garde, and G. Hummer, *Ann. Rev. Phys. Chem.* **59**, 713 (2008).
  - [54] S. Vembanur, A. J. Patel, S. Sarupria, and S. Garde, *J. Phys. Chem. B* **117**, 10261 (2013).
  - [55] J. D. Weeks, D. Chandler, and H. C. Andersen, *J. Chem. Phys.* **54**, 5237 (1971).
  - [56] R. Underwood and D. Ben-Amotz, *J. Chem. Phys.* **135**, 201102 (2011).
  - [57] N. Choudhury and B. M. Pettitt, *J. Am. Chem. Soc.* **127**, 3556 (2005).
  - [58] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
  - [59] "See supplemental material at <http://dx.doi.org/10.1063/1.4905009>."
  - [60] H. C. Andersen, J. D. Weeks, and D. Chandler, *Phys. Rev. A* **4**, 1597 (1971).
  - [61] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
  - [62] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *J. Chem. Theory Comp.*, 435 (2008).
  - [63] G. Bussi, D. Donadio, and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).
  - [64] M. Parrinello and A. Rahman, *J. Applied Phys.* **52**, 7182 (1981).
  - [65] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995).
  - [66] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford: New York, 1987).
  - [67] Z. Tan, E. Gallichio, M. Lapelosa, and R. M. Levy, *J. Chem. Phys.* **136**, 144102 (2012).
  - [68] M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**, 124105 (2008).
  - [69] G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
  - [70] T. T. Pham and M. R. Shirts, *J. Chem. Phys.* **136**, 124120 (2012).
  - [71] L. N. Naden, T. T. Pham, and M. R. Shirts, *J. Chem. Theory Comput.* **10**, 1128 (2014).
  - [72] We estimate an upper bound for the magnitude of solute-water attractions to be  $\lambda_i \beta \epsilon \approx 0.7$ . By comparison, a square-well potential for the  $R_C = 0.6$  nm solute, with

a well-depth,  $\lambda_i\beta\epsilon \approx 0.71$ , has the same second virial co-efficient as the highly attractive C<sub>60</sub>-water interaction potential.

- [73] L. Maibaum and D. Chandler, J. Phys. Chem. B **111**, 9025 (2007).
- [74] Rahul, S. N. Jamadagni, V. Venkateshwaran, and S. Garde, “Connecting water correlations, fluctuations, and wetting phenomena at hydrophobic and hydrophilic surfaces,” arxiv:1409.2570 (2014).
- [75] A. P. Willard and D. Chandler, J. Chem. Phys. **141**, 18C519 (2014).
- [76] N. Galamba, J. Phys. Chem. B **117**, 2153 (2013).
- [77] J. D. Weeks, K. Katsov, and K. Vollmayr, Phys. Rev. Lett. **81**, 4400 (1998).
- [78] Y. G. Chen, C. Kaur, and J. D. Weeks, J. Phys. Chem. B **108**, 19874 (2004).
- [79] Y. G. Chen and J. D. Weeks, Proc. Natl. Acad. Sci. USA **103**, 7560 (2006).
- [80] J. M. Rodgers and J. D. Weeks, J. Phys.: Condens. Matter **20**, 494206 (2008).
- [81] J. M. Rodgers and J. D. Weeks, Proc. Natl. Acad. Sci. USA **105**, 19136 (2008).
- [82] R. C. Remsing, J. M. Rodgers, and J. D. Weeks, J. Stat. Phys. **145**, 313 (2011).
- [83] R. C. Remsing, *From Structure to Thermodynamics with Local Molecular Field Theory*, Ph.D. thesis, University of Maryland (2013).
- [84] Y. Shi and T. L. Beck, J. Chem. Phys. **139**, 044504 (2013).
- [85] K. D. Collins, Biophys. Chem. **167**, 43 (2012).

# Supplementary Material for “Water Density Fluctuations Relevant to Hydrophobic Hydration are Unaltered by Attractions”

Richard C. Remsing<sup>2</sup> and Amish J. Patel<sup>2,\*</sup>

<sup>2</sup>*Department of Chemical and Biomolecular Engineering,  
University of Pennsylvania, Philadelphia, PA 19104*

(Dated: February 19, 2015)

## II. DERIVATION OF EQUATIONS 5 AND 6 OF THE MAIN TEXT

The probability distribution of the number of waters in the core and shell volumes can be written as

$$P_{v,v_{\text{sh}}}^{(i)}(N, N_{\text{sh}}) = \langle \delta_{N,N_C} \delta_{N_{\text{sh}},N_S} \rangle_{\lambda_i} \quad (\text{S1})$$

$$= \frac{Z_0}{Z_i} \left\langle \delta_{N,N_C} \delta_{N_{\text{sh}},N_S} e^{-\beta \lambda_i U(N_C, N_S)} \right\rangle_0, \quad (\text{S2})$$

where  $Z_0$  and  $Z_i$  are the partition functions in the absence and presence of the attractive potential, respectively. The delta functions allow the operators  $N_C$  and  $N_S$  to be replaced by the corresponding variables  $N$  and  $N_{\text{sh}}$ , so that

$$P_{v,v_{\text{sh}}}^{(i)}(N, N_{\text{sh}}) = C \langle \delta_{N,N_C} \delta_{N_{\text{sh}},N_S} \rangle_0 e^{-\beta U(N, N_{\text{sh}})} \quad (\text{S3})$$

$$= C P_{v,v_{\text{sh}}}(N, N_{\text{sh}}) e^{-\beta U(N, N_{\text{sh}})}, \quad (\text{S4})$$

where  $C \equiv Z_0/Z_i$ . Equation 5 of the main text is then obtained by summation over  $N_{\text{sh}}$ .

The free energy  $\Delta G_i$  of turning on an attractive potential in bulk water is

$$e^{-\beta \Delta G_i} = \frac{Z_i}{Z_0} = \left\langle e^{-\beta \lambda_i U(N_C, N_S)} \right\rangle_0. \quad (\text{S5})$$

By inserting the identity twice, we obtain

$$e^{-\beta \Delta G_i} = \left\langle \sum_N \sum_{N_{\text{sh}}} \delta_{N,N_v} \delta_{N_{\text{sh}},N_{v_{\text{sh}}}} e^{-\beta \lambda_i U(N_v, N_{v_{\text{sh}}})} \right\rangle_0 \quad (\text{S6})$$

$$= \sum_N \sum_{N_{\text{sh}}} \left\langle \delta_{N,N_v} \delta_{N_{\text{sh}},N_{v_{\text{sh}}}} \right\rangle_0 e^{-\beta \lambda_i U(N, N_{\text{sh}})} \quad (\text{S7})$$

$$= \sum_N \sum_{N_{\text{sh}}} P_{v,v_{\text{sh}}}(N, N_{\text{sh}}) e^{-\beta \lambda_i U(N, N_{\text{sh}})}, \quad (\text{S8})$$

resulting in Equation 6 of the main text.

## III. DENSITY FLUCTUATIONS IN THE PRESENCE OF ATTRACTIONS IN SMALL AND LARGE VOLUMES

In Figure S1 and Figure S2, we show the distributions  $P_v^{(1)}(N)$  and  $P_v^{(2)}(N)$  for the methane sized solute and

a solute with  $R_C = 0.9$  nm, respectively. In both cases, the presence of the attractive potential,  $\lambda_1 U_1$ , decreases the probability of the low  $N$  tails of the distribution, independent of whether the unperturbed distribution is Gaussian, as in Figure S1a, or non-Gaussian, as in Figure S2a. In contrast, water density fluctuations around the mean are unaltered by the presence of attractions in the hydration shell alone,  $\lambda_2 U_2$ , and the distributions obtained for both the methane-sized volume (Figure S1b) and the large  $R_C = 0.9$  nm volume (Figure S2b) only exhibit small decreases in  $\langle N_v(\mathbf{R}) \rangle_{\lambda_2}$  with increasing  $\lambda_2$ . These findings support the conclusions drawn in the main text.

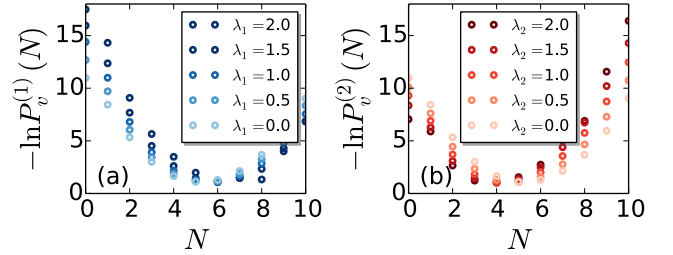


FIG. S1: Density fluctuations in the presence of the attractive fields (a)  $\lambda_1 U_1$  and (b)  $\lambda_2 U_2$  for a spherical solute with  $R_C = 0.336$  nm and  $R_S = 0.636$  nm.

## IV. LENNARD-JONES ATTRACTIONS

Here, we demonstrate that the conclusions drawn in the main text for a square-well potential are generally

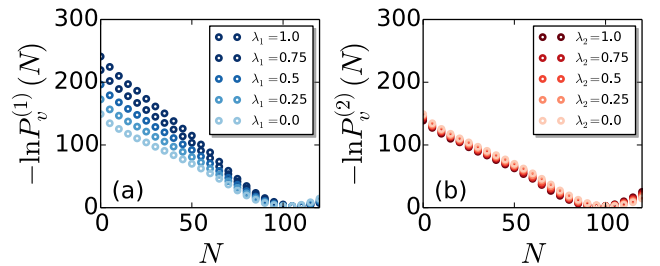


FIG. S2: Density fluctuations in the presence of the attractive fields (a)  $\lambda_1 U_1$  and (b)  $\lambda_2 U_2$  for a spherical solute with  $R_C = 0.9$  nm and  $R_S = 1.2$  nm.

\* amish.patel@seas.upenn.edu

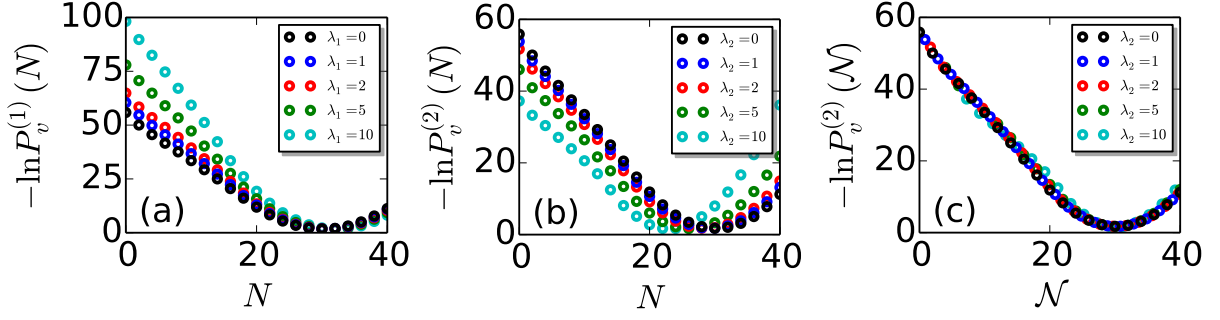


FIG. S3: Water number fluctuations,  $-\ln P_v^{(i)}(N)$ , in the core region of a solute with radius,  $R_C = 0.6$  nm, in the presence of attractive fields corresponding to the (a) full LJ attractions and (b) LJ attractions acting in the hydration shell alone. (c) The distributions corresponding to attractions in the shell alone, are shifted horizontally to align their means, highlighting that density fluctuations are not modified by LJ attractions in the hydration shell.

applicable. Following Huang and Chandler [23], we use the integrated 9–3 LJ potential, separated into purely repulsive and purely attractive components,  $u_0^{\text{LJ}}(r)$  and  $u_1^{\text{LJ}}(r)$ , respectively, following the WCA prescription [55]. We then linearly couple the attractive potential to the parameter  $\lambda_1$  in order to tune the strength of the attractive interactions, and study how the application of the corresponding attractive potential,  $\lambda_1 U_1^{\text{LJ}}$ , alters density fluctuations in the bulk. The LJ-like analog of the  $U_2$  potential is then obtained by applying an additional potential,  $U_3$ , which subtracts the attractions imposed by  $U_1^{\text{LJ}}$  in the solute core, as follows:

$$U_2^{\text{LJ}}(\mathbf{R}) = U_1^{\text{LJ}}(\mathbf{R}) + U_3(\mathbf{R}) = U_1^{\text{LJ}}(\mathbf{R}) - u_{\min} N_v(\mathbf{R}), \quad (\text{S1})$$

where  $u_{\min}$  is the minimum of the solute-solvent pair potential. We then linearly couple this potential to a parameter  $\lambda_2$  in order to vary the magnitude of the attractive interactions. The core region,  $v$ , is now defined by the effective hard core radius of the solute, which is estimated as [23, 60].

$$R_C \approx \int_0^\infty dr \{1 - \exp[-\beta u_0^{\text{LJ}}(r)]\}. \quad (\text{S2})$$

The probability  $P_v^{(2)}(N)$  in the presence of this attractive potential,  $\lambda_2 U_2^{\text{LJ}}$ , is then determined through reweighting the corresponding distribution obtained in the presence of  $\lambda_1 U_1^{\text{LJ}}$ . Thus, for  $\lambda_2 = \lambda_1$ ,

$$P_v^{(2)}(N) = C P_v^{(1)}(N) e^{\beta \lambda_1 u_{\min} N}, \quad (\text{S3})$$

where  $C$  is a normalization constant. Density fluctuations in the absence and presence of the attractive potentials in the solute core are shown in Figures S3a and S3b, respectively, for a solute with  $R_C = 0.6$  nm, and display the same qualitative trends as the square well potential.

## V. DERIVATION OF THE LINEAR RESPONSE APPROXIMATION LR3

In order to illustrate the connection of LR3 to the other linear response approximations, we start with the exact expression for the free energy for turning on the potential,  $\lambda_i U_i(\mathbf{R})$ , in bulk water:

$$\beta \Delta G_i = -\ln \left\langle e^{-\beta \lambda_i U_i(\mathbf{R})} \right\rangle_0, \quad (\text{S1})$$

where  $\langle \dots \rangle_0$  is the ensemble average over configurations  $\mathbf{R}$  obtained in the absence of the attractive potential. The linear response approximation LR2 is obtained from a second order cumulant expansion of the right hand side of Equation S1,

$$\beta \Delta G_i \approx \beta \lambda_i \langle U_i(\mathbf{R}) \rangle_0 - \frac{\beta^2 \lambda_i^2}{2} \langle (\delta U_i(\mathbf{R}))^2 \rangle_0. \quad (\text{S2})$$

Because the potentials considered here have the form,  $U_i(\mathbf{R}) = \epsilon N_{v_i}(\mathbf{R})$ , the free energy becomes

$$\beta \Delta G_i \approx \lambda_i \beta \epsilon \langle N_{v_i}(\mathbf{R}) \rangle_0 - \frac{(\lambda_i \beta \epsilon)^2}{2} \langle (\delta N_{v_i}(\mathbf{R}))^2 \rangle_0. \quad (\text{S3})$$

We can now use the macroscopic relation between particle number fluctuations and the isothermal compressibility,  $\langle (\delta N_{v_i}(\mathbf{R}))^2 \rangle_0 = \rho_B k_B T \kappa_T \langle N_{v_i}(\mathbf{R}) \rangle_0$ , in conjunction with  $\langle N_{v_i}(\mathbf{R}) \rangle_0 = \rho_B v_i$ , to obtain our final expression for LR3,

$$\Delta G_i \approx \lambda_i \epsilon \rho_B v_i - \frac{(\lambda_i \epsilon \rho_B)^2}{2} \kappa_T v_i, \quad (\text{S4})$$

where  $v_i$  is the volume over which the potential is applied. The use of macroscopic compressibility in Equation S4 is rigorously true only for large (infinite) volumes. Equation S4 is specific to our model square well attraction potentials; a general form applicable to all slowly-varying potentials is derived in the next section.



## VI. LINEAR RESPONSE THEORY FOR TURNING ON SLOWLY-VARYING POTENTIALS

The derivation of the generalized analog of LR3 follows the work of Remsing and Weeks [83] for charged systems. For generality, consider turning on an attractive potential  $w_1^{(\lambda)}(\mathbf{r})$  in a uniform fluid, where  $w_1^{(\lambda)}(\mathbf{r})$  is a slowly-varying attractive potential coupled to a parameter  $\lambda$  in some fashion that is not necessarily linear. Within the linear response regime, the free energy of turning on the attractive potential is given by

$$\Delta G_\lambda \approx \frac{1}{2} \left[ \langle W_1^{(\lambda)}(\bar{\mathbf{R}}) \rangle_0 + \langle W_1^{(\lambda)}(\bar{\mathbf{R}}) \rangle_\lambda \right], \quad (\text{S1})$$

where  $W_1^{(\lambda)}(\bar{\mathbf{R}}) = \sum_{i=1}^{N_{\text{tot}}} w_1^{(\lambda)}(\mathbf{r}_i; \bar{\mathbf{R}})$  is the total energy due to the interaction of the  $N_{\text{tot}}$  particles in the system with the attractive potential. Such an approximation has been shown to be exceptionally accurate when  $w_1^{(\lambda)}(\mathbf{r})$  is *slowly-varying* over typical correlation lengths in the fluid [83].

$\Delta G_\lambda$  can be equivalently written in terms of the non-uniform singlet density,

$$\Delta G_\lambda = \frac{1}{2} \int d\mathbf{r} [\rho_\lambda(\mathbf{r}) + \rho_B] w_1^{(\lambda)}(\mathbf{r}), \quad (\text{S2})$$

where  $\rho_\lambda(\mathbf{r}) = \langle \rho(\mathbf{r}; \bar{\mathbf{R}}) \rangle_\lambda = \langle \sum_{i=1}^{N_{\text{tot}}} \delta(\mathbf{r} - \mathbf{r}_i(\bar{\mathbf{R}})) \rangle_\lambda$  is the average non-uniform density in the presence of the field  $w_1^{(\lambda)}(\mathbf{r})$ , and  $\rho_B$  is bulk density.

To approximate  $\rho_\lambda(\mathbf{r})$ , we perform a functional expansion of the density, such that the deviation from the bulk density,  $\Delta\rho_\lambda(\mathbf{r}) \equiv \rho_\lambda(\mathbf{r}) - \rho_B$  is

$$\Delta\rho_\lambda(\mathbf{r}) \approx -\beta \int d\mathbf{r}' \chi_0(|\mathbf{r} - \mathbf{r}'|) w_1^{(\lambda)}(\mathbf{r}'), \quad (\text{S3})$$

which, in Fourier space, is equivalent to

$$\Delta\hat{\rho}_\lambda(\mathbf{k}) \approx -\beta \hat{\chi}_0(\mathbf{k}) \hat{w}_1^{(\lambda)}(\mathbf{k}). \quad (\text{S4})$$

Here,  $\chi_0(\mathbf{r}, \mathbf{r}') = \langle \delta\rho_0(\mathbf{r}; \bar{\mathbf{R}}) \delta\rho_0(\mathbf{r}'; \bar{\mathbf{R}}) \rangle_0$  is the bulk density-density correlation function, and  $\delta\rho_0(\mathbf{r}; \bar{\mathbf{R}}) = \rho(\mathbf{r}; \bar{\mathbf{R}}) - \rho_B$ .

Because of the slowly-varying nature of  $w_1^{(\lambda)}(\mathbf{r})$ , we expect only small values of  $k$  to contribute to the density response and therefore the free energy. In the limit of small  $k$ , the linear response function is

$$\hat{\chi}_0(\mathbf{k}) \sim \beta^{-1} \rho_B^2 \kappa_T, \quad (\text{S5})$$

where  $\kappa_T$  is the isothermal compressibility of the fluid. The density response can then be approximated by

$$\Delta\hat{\rho}_\lambda(\mathbf{k}) \approx -\beta \hat{\chi}_0^{(0)} \hat{w}_1^{(\lambda)}(\mathbf{k}) = -\rho_B^2 \kappa_T \hat{w}_1^{(\lambda)}(\mathbf{k}), \quad (\text{S6})$$

to zeroth order in  $k$ . Thus, in real space, the induced density response is

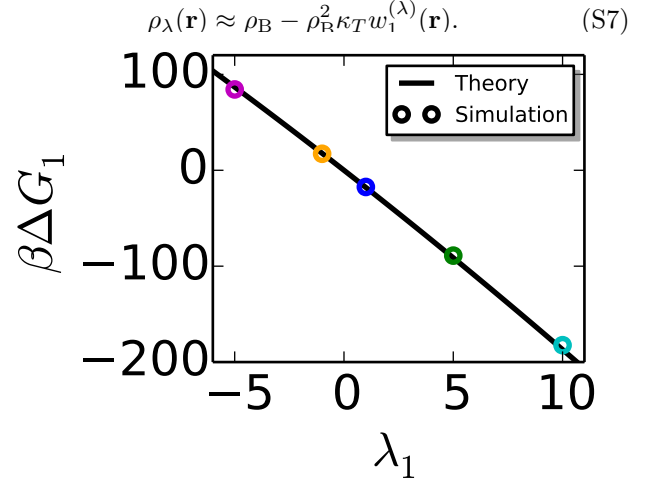


FIG. S4: The free energy of turning on the 9-3 LJ attractive potential  $\lambda_1 u_1^{\text{LJ}}(r)$  can be accurately described by the linear response theory Equation S8, as shown here for the  $R_C = 0.6$  nm spherical solute.

We can now use this approximation for the density response to determine the free energy, yielding

$$\Delta G_\lambda \approx -2a_\lambda \rho_B - \frac{\rho_B^2 \kappa_T}{2} \int d\mathbf{r} \left( w_1^{(\lambda)}(\mathbf{r}) \right)^2, \quad (\text{S8})$$

where

$$a_\lambda \equiv -\frac{1}{2} \int d\mathbf{r} w_1^{(\lambda)}(\mathbf{r}), \quad (\text{S9})$$

in analogy with the van der Waals constant.

We test these linear response theory predictions for the free energy of turning on the integrated LJ attractive potential  $u_1^{\text{LJ}}(r)$  in the bulk solvent. Specifically, Figure S4 shows the free energy of turning on  $\lambda_1 u_1^{\text{LJ}}(r)$  for a solute with radius  $R_C = 0.6$  nm, both from simulation and from Equation S8. The theoretical predictions are in very good agreement with simulation results; the slowly-varying nature of the LJ potential makes it even more amenable to treatment with linear response theory than the square-well potential studied in the main text. Additionally, because water is highly incompressible (small  $\kappa_T$ ), we find that for typical values of  $\lambda_1$ , the term linear in  $\lambda_1$  in Equation S8 dominates.